

ユーザの移動を考慮した 多対多型マルチキャスト配送木の動的構成手法

福田 光利[†] 中村 豊^{††} 藤川 和利^{††} 砂原 秀樹^{††}

[†] 奈良先端科学技術大学院大学 情報科学研究科 〒630-0192 奈良県生駒市高山町 8916-5

^{††} 奈良先端科学技術大学院大学 情報科学センター 〒630-0192 奈良県生駒市高山町 8916-5

E-mail: [†]mitsu-fu@is.aist-nara.ac.jp, ^{††}{yutaka-n,fujikawa,suna}@itc.aist-nara.ac.jp

あらまし ネットワークゲームでは多数のユーザ同士のデータ交換のため多対多型のマルチキャスト通信を利用することが考えられる。この場合ゲーム空間でのユーザ同士の関係に基づいてマルチキャスト通信を行うため Application Layer Multicast(ALM) が用いられる。従来のマルチキャストではユーザ端末が移動することが想定されていないため、移動端末を用いるユーザが移動し任意の場所からインターネットに接続するような場合 ALM の配送木が必ずしも最適な配送経路をたどるとは限らない。また多対多型で配信する場合でも全てのユーザに配信する必要があるとは限らないため複数の ALM を用いることが望まれ、ゲーム空間全体で複数の ALM のメンバ管理を動的に管理する必要がある。そのため本研究では端末の移動に伴い動的に配送木を再構成する ALM 方式を提案する。また、複数の ALM 間でメンバ情報を交換し適切なマルチキャストグループを形成する機構も提案する。これにより多人数参加型のアプリケーション上でユーザ同士のデータ交換を効率よく行うことが可能となる。

キーワード Application Layer Multicast, グループコミュニケーション, メンバ管理

Dynamic Topology Configuration System of N-toN Multicast Delivery Tree for Mobile User

Mitsutoshi FUKUDA[†], Yutaka NAKAMURA^{††}, Kazutoshi FUJIKAWA^{††},

and Hideki SUNAHARA^{††}

[†] Graduate School of Information Science of NARA Institute of Science and Technology Takayama-cho
8916-5, Ikoma-shi, Nara, 630-0192 Japan.

^{††} Information Technology Center of NARA Institute of Science and Technology Takayama-cho
8916-5, Ikoma-shi, Nara, 630-0192 Japan

E-mail: [†]mitsu-fu@is.aist-nara.ac.jp, ^{††}{yutaka-n,fujikawa,suna}@itc.aist-nara.ac.jp

Abstract In an online network game, it is desirable to use N-to-N multicasting communication for many users' data exchange. In this case, in order to perform multicasting communication based on the relation of the users in game space, Application Layer Multicast (ALM) is used. In the conventional multicasting, since it is not assumed that a user terminal migrate. In this paper, the system of ALM which reconstructs a delivery tree dynamically with movement of a terminal is proposed. Moreover, the mechanism which exchanges member information among two or more ALM, and forms a suitable multicasting group is also proposed.

Key words Application Layer Multicast, group communication, node management

1. はじめに

現在のインターネットでは、複数のユーザ間でグループを形成し、その中で情報を交換する形態のアプリケーションが重要な位置を占めている。本研究では、このような形態のコミュニ

ケーションをグループコミュニケーションと呼ぶ。グループコミュニケーションにはチャットシステムやテレビ会議、ネットワークゲームなどがある。近年、ホットスポットなどでモバイル端末を利用する環境が整いつつある。今後グループコミュニケーションに参加するノードは有線・無線混在していくと考え

られる。またモバイル端末を利用しているユーザは移動することが考えられる。多地点間で多対多のグループコミュニケーションを行う上で重要な技術となるのがマルチキャスト技術である。マルチキャストは送信されたデータが中継ノードによって必要に応じて複製され、複数の受信者に送信されるため複数の宛先に対する効率的な同報通信を実現する。送信者は一つのデータを送信するだけでよい。現行のマルチキャストはネットワーク層でマルチキャスト配送木を構築する IP Multicast とネットワーク層でマルチキャストを実現する Application Layer Multicast(ALM)[1][2][3][4][5][6]がある。

IP Multicast は大人数に対して効率的な配信が可能である。また IP Multicast は現在ほとんどのルータに実装されている。IP Multicast では一般ユーザは全てマルチキャスト配送木の末端であり途中で離脱しても他のユーザのコミュニケーションに影響はないという特徴がある。

しかし IP Multicast にはいくつかの問題点がある。IP Multicast ではネットワーク層におけるプロトコルが複雑になり、その上で実装されるアプリケーションは多くの制約を受ける。そのため、アプリケーションの自由度は高くない。

またインター・ドメイン・ルーティングの困難さ、マルチキャストアドレス割当の困難さ、アクセス制御やグループ管理の困難さという問題があり、このような理由からインターネットへの展開は進んでいない。また有線ノード・無線ノードが混在している環境での IP Multicast の実装は困難である。図1に IP Multicast の例を示す。

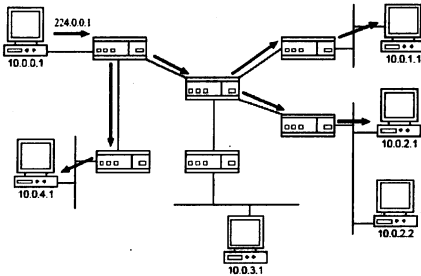


図1 IP Multicast の例

ALM ではデータはユニキャストを用いてノード間をピア・ツー・ピアで中継し、最終的に全ノードに伝送される。つまり、ノードだけを構成要素とするオーバーレイネットワークの上にマルチキャストツリーを形成し、そのツリーに沿ってノード間をパケットリレー式にデータを伝送する。従って有線ノード・無線ノードが混在していてもネットワーク層に左右されずマルチキャスト配送木を構成できる。

ALM はマルチキャストに関する機能、即ちパケットの複製やマルチキャストルーティング、グループ管理などをアプリケーション層で実現するためネットワーク層はユニキャストのみをサポートしていればよい。しかし、アクセス制御やグループ管理の困難さそのものは IP Multicast と変わらない。また ALM ではエンドノードが中継ノードとなりマルチキャスト配送木を

構成するため中継ノードとなっているユーザが移動や離脱をした場合マルチキャスト配送木が切れるという問題がある。

本研究ではネットワーク層でマルチキャストルータが普及していない環境においても利用可能である ALM 技術を用いて一般ユーザが移動する環境において多対多型のグループコミュニケーションの実現のためマルチキャスト配送木の構成を動的に行う仕組みの提案をする。

2. マルチキャスト配送方式

2.1 マルチキャストを行う配送方式

本節ではマルチキャストを行う配送方式を述べる。

以下にマルチキャスト経路制御方式に利用可能な方式として Flooding, Spanning Tree, Source Based Tree (SBT), Shared Tree を挙げる。

(1) 単純な方法

- Flooding
- Spanning Tree

Flooding と Spanning Tree は原始的なマルチキャスト経路制御方式を構築可能なアルゴリズムである。これらの原始的なアルゴリズムをそのまま用いると帯域や関係するマルチキャストルータの資源を多く消費する。

(2) Source Based Tree (SBT)

- Reverse Path Broadcasting(RPB)

全てのネットワーク間に1つの Spanning Tree を構築するよりより効果的な解決策は各サブネットワークに対してマルチキャストグループごとに Spanning Tree を構築することである。これらの Spanning Tree はサブネットワークが直接送信元とつながっていることから結果として送信側ベースの配送木を構築することになる。送り先が多数存在する可能性があるので多くの配送木を構築しなくてはならない。

- Truncated Reverse Path Broadcasting(TRPB)

TRPB は RPB の欠点を克服するために発展してきた。IGMP の助けによりマルチキャストルータは各サブネットワーク上のグループメンバーシップを決定し少なくとも1人のグループメンバーがそのサブネットワーク上にいなければデータグラムを送らないようにできる。従って葉サブネットワークに誰もグループメンバーがいなければ配送木はルータによって刈りとられる。

- Reverse Path Multicast(RPM)

RPM はグループメンバーのいるサブネットワーク、グループメンバーのいるサブネットワークへの最短の経路沿いのルータやサブネットワークのときのみ配送木を構築する。

(3) Shared Tree 方式

Shortest path tree アルゴリズムのように、各 (Src, Group) 毎に異なる配送木を構築するのではなくグループのメンバー全員で共有される1つの配送木を構築する。Shared

Tree アルゴリズムは Spanning Tree アルゴリズムに非常に似ているがグループごとに違った配送木を構築できるといふ点異なる。マルチキャストのパケットを受けとりた場合には、Shared delivery tree に参加しなくてはならない。送信元に関係なく同じ配送木でマルチキャストのパケットは送られる。

2.2 マルチキャスト配送方式の比較

本節ではデータ配送方式、配送木の維持管理コスト、多対多型通信対応、耐障害性、既存の ALM の比較を行う。

2.2.1 データ配送方式の比較

マルチキャストのメッセージ配送方式は、一般的に表 1 のように分類される [7]。Flooding 方式はブロードキャストとして実装されるため全く効率化されていない配送方式である。また Spanning Tree 方式も Flooding 方式と同じくグループという概念がないためネットワーク全体にツリーが構成されてしまう。しかしツリーの形状そのものはシンプルでかつループの発生がないという利点がある。

SBT 方式では送信者毎にルーティング情報が必要なためルーティングテーブルが大きくなる問題や、ツリーを再構築する際のフラッディングの問題がある。しかし、各送信元に最適化されたツリーが構築できる利点がある。

Shared Tree 方式は Rendezvous Point(RP) の決定が困難であるが、共有ツリーは経路情報が少なく効率的なグループ管理が可能である。

表 1 データ配送方式の比較

配信方式	実現の方法	利点	問題点
Flooding	重複メッセージは放棄するが、それ以外は送信先に送った全ての隣接ノードへ送信	ルーティングのために隣接ノード以外の情報を持つ必要が無い	ブロードキャストのためトラフィックが増加。メッセージの重複が発生
Spanning Tree	ルーティング先はメッセージの送信元以外のツリーに送った全ての隣接ノードとなる	単純な構造で、かつループの発生がない	ルーティングの中間位置のノードにトラフィックが集中することがある
Source Based Tree	ルーティング先はメッセージの送信元毎に指定された隣接ノードとなる	各送信元ノード毎に最適化されたルーティング経路を持つことができる	送信元毎に経路情報が必要
Shared Tree	送信元ノードは RP に対して送信する。ルーティング先は全ノードで共有する配信ツリーに指定された隣接ノードとなる	ツリーを共有するため経路情報が少ない	最適なルーティング経路ではない。RP が必要で負荷が集中する

2.2.2 動的なマルチキャスト配送木の維持

表 2 にそれぞれの配信方式においてマルチキャストグループからの離脱や、配送木の再構築が生じた場合の影響についてまとめる。

Flooding 方式、Spanning Tree 方式ではノードの離脱・移動に対する配送木再構築のコストが比較的小さい。

2.2.3 多対多型通信への対応

表 3 は全てのグループに属するユーザが送信者になることを考慮した上で比較したものである。Flooding 方式と SBT 方式は多対多型通信方式のマルチキャストに適しているといえる。

2.2.4 耐障害性の比較

ユーザが故障により離脱した場合における各方式の比較をする。表 4 は各方式における耐障害性についてまとめたものであ

表 2 配送木の維持管理コストの比較

配信方式	影響範囲	備考	
Flooding	ネットワーク再構築によって隣接関係に変化があったノード	隣接ノードに従って情報を更新する	影響範囲小
Spanning Tree	ネットワーク再構築によって隣接関係に変化があったノード	新たな Spanning Tree 上の隣接ノードに従ってルーティング情報を更新する	影響範囲小
Source Based Tree	ネットワークの再構築によって一部でも影響のある配信ツリーに属する全てのノード	影響ノードに対して送信元ツリー全ての再構築処理を行う必要がある	再構築時にフラッディングが発生する。送信元に経路情報の更新が必要
Shared Tree	ネットワークの再構築によって隣接ノードに変化があったノード	新たな共有ツリー上の隣接ノードをルーティング情報によって更新	RP が離脱した場合ツリー全体の再構築が必要

表 3 多対多型通信対応での比較

配信方式	影響範囲	備考	適用の可否
Flooding	配信ツリーが構築されないため送信元に依存しない	ネットワーク全体にメッセージがフラッディングされるため	○ (送信元に依存しない)
Spanning Tree	配信ツリーの形状に依存する	MST (Minimum Spanning Tree) となっていれば最適	△ (ツリーの形状に依存)
Source Based Tree	全ての配信元に最適なツリーを構築できる	全ての配信元に応じた経路情報が必要	○ (全ての送信元に最適なツリーが形成される)
Shared Tree	配信ツリーの形状に依存する	全てのメッセージは RP を経由する	△ (ツリーの形状に依存)

る。表のように Flooding 方式や SBT 方式は、ほとんど影響を受けないが多くの 2 方式、特に Shared Tree 方式では RP のダウン時の影響が大きいといえる。

表 4 耐障害性の比較

配信方式	影響	備考	適用の可否
Flooding	なし	次回配信時にルーティング情報を更新	○ (影響小)
Spanning Tree	分断される	新たな隣接メンバを探索する必要がある	△ (隣接メンバの探索メカニズムが必要)
Source Based Tree	なし	次回再構築時にルーティング情報を更新	○ (影響小)
Shared Tree	分断されてしまう	RP に障害が発生してしまうと影響が大きい	× (影響大)

2.2.5 既存の ALM の比較

ALM にはいきなり配送木を構築する Tree-first のアプローチと、メッシュ状のオーバーレイネットワークを構築してから配送木を構築する Mesh-first のアプローチの 2 つに大きく分類される。

Tree-first のアプローチは Mesh-first のアプローチよりも経路の冗長性の低さから耐障害性が低い。既存の ALM の方式を表 5 にまとめた。

表 5 既存の ALM の比較

アプローチ	Tree-first	Mesh-first
ツリーの形状	Shared Tree	Source Based Tree
ツリーの最適化	管理サーバ方式、分散強調方式 (必ずしも最適ではない)	送信時に最適なツリーを形成
経路の冗長性	小	中
方式例	ALMI [2], Overcast [4], Yoid [3]	Narada [1], Gossamer [5]

2.3 考察

以上より、マルチキャスト配送木構成方式について比較を行っ

た。まず Flooding 方式は経路が最適化されずトラフィックも増加し移動端末の存在する ALM への適用は不適切である。SBT 方式は経路は最適になるが、配送木構築の際の Flooding が問題となる。Spanning Tree 方式はネットワーク全体に配送木が構築されるためグループ毎に配送木を構築はできない。Shared Tree 方式は RP の決定が問題となる。本研究では Shared Tree 方式を取り入れマルチキャスト配送木を構成する際の RP 決定に関して提案を行う。

2.4 多対多型コミュニケーションの概要

多地点間における多対多型のコミュニケーションとは多対多型の通信形態をとる。代表的なものとしてチャットシステムや、テレビ会議システムネットワークゲームなどが挙げられる。特定のグループの中で情報を共有するために用いられ、双方向のコミュニケーションである場合が多い。情報を共有したい複数のユーザがグループを形成し、グループ内の全てのユーザに対して情報を送信する。また全てのユーザから情報を受信することもできる。

2.4.1 多対多型通信システムのアーキテクチャ

多地点間で多対多のグループコミュニケーションを行うシステムには、システムの目的やネットワーク環境に応じたいくつかのアーキテクチャが存在する。以下に代表的なアーキテクチャであるサーバレス方式、機能集中型サーバ方式、一部ピア・ツー・ピア方式、機能分散型サーバ方式について述べる。

(1) サーバレス方式

メンバ管理に専用のサーバを置かない方式で、アーキテクチャ上最もシンプルなアーキテクチャである。各ホストは他のホストへの更新情報を通知する機能を持たせている。このアーキテクチャは各ホストやネットワークに高い性能を求めるもので、スケーラビリティにも問題がある。

(2) 機能集中型サーバ方式

メンバ情報のデータを保持する専用のサーバを一つ設置し、一般ユーザがメンバ管理サーバと通信して配送木を構築するアーキテクチャである。サーバと一般ユーザ間の通信処理部分が比較的シンプルであり、機能に応じたサーバを選択する必要がないため一般ユーザ側の負荷が少ない。しかしサーバ側に負荷が集中し、負荷分散の形態についてバリエーションが少ないという欠点がある。

(3) 一部ピア・ツー・ピア方式

(2)に加えて、必要な機能のみ一般ユーザ間のピア・ツー・ピア接続方式で対処をするアーキテクチャである。多人数参加型のサービスを行う場合は(1)と同様の課題が存在する。またメンバ管理サーバは一つであるため(2)で挙げた問題も含む。

(4) 機能分散型サーバ方式

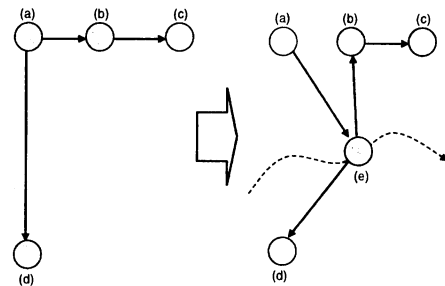
機能やサービスに応じてサーバを分ける方式で、機能の追加や負荷の分散形態に柔軟に対応できる。一般ユーザはその時々に必要なサーバとのみ通信をすれば良く、一般ユーザの処理能力に応じて選択することができる。

2.4.2 システムに求められる課題

本研究で求めるシステムでは移動する一般ユーザへの対応、スケーラビリティ、中継ノードのダウンへの対応が求められる。

(1) 移動する一般ユーザへの対応

グループコミュニケーションに参加している一般ユーザの物理的な位置の移動が発生し、利用しているネットワークインフラストラクチャに変化が発生した場合、配送木の変更を行う。配送木を変更する基準や手順をどのように行うかが挙げられる。図2では既存のノード(e)が移動してきたことにより既存の配送木が変化した例を示している。図の左側の配送木ではノード(a)が配送木のコアとなっているが右側の配送木ではノード(e)が配送木のコアとなっている。



移動するノード(e)によって変化する配送木

図2

(2) スケーラビリティ

参加しているユーザが増加しても少人数の場合と同様のサービスを提供できる性能が求められる。そのような性能を持つシステムは、スケーラビリティがあるシステムと呼ばれる。スケーラビリティは、ある処理に対するシステムの応答性が低下する場合に損なわれる。その原因としてユーザの増加によりある処理が完了するまでの時間が増加することが挙げられる。

(3) 中継ノードのダウンへの対応

ALM 技術をマルチキャストデータグラム配送を行う際の問題点の一つとして、中継ノードが不意に切断された場合の対応がある。ALM では一般ユーザが中継ノードとなり得るのでグループからの離脱要求をメンバ管理ノードに送れずに不意に切断される事がある。そのような状況になった時、中継ノードの下流にいる一般ユーザへの到達性が失われる問題がある。これを図3に示す。

2.5 システムの機能要件

2.5.1 想定環境

対象とするグループコミュニケーションはオンラインゲーム

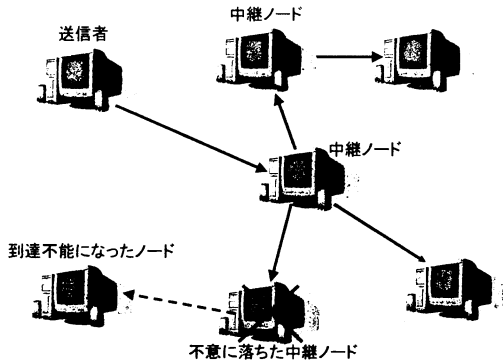


図3 中継ノードが落ちた例

などを想定しておりグループを構成する要件は以下の通りとする。

- グループの規模:メンバ数 100~1000 程度
- メッセージの形式:特にこだわらない
- メッセージの転送:ALM によって行う
- グループの数:数十程度

2.5.2 システムを実現するために必要な処理

グループコミュニケーションを実現するためには以下の4つの処理が必要となる。

- (1) マルチキャストホストグループが定義される必要がある。ホストグループは、マルチキャストアドレスや URI を識別子として取り扱われている。
- (2) 一般ユーザがマルチキャストグループに参加、また離脱する際のメカニズムが必要となる。これには、送信側または受信側ベースのグループメンバシップに関する情報を、ネットワーク全体に伝送し、管理するためのプロトコルも含まれる。
- (3) マルチキャストルーティングプロトコルが、必要に応じてマルチキャストトラフィックのコピーや、転送を行う必要がある。
- (4) マルチキャストセッションで配信されるデータの作成と、管理を行うアプリケーションプロトコルが必要となる。

2.5.3 システムに求められる要素と機能

システムを構成するのに必要な要素とその機能を以下に示す。

- (1) サービス提供者
 - サービスの提供を行う
- (2) メンバ管理ノード
 - グループへ参加の受付・処理
 - グループから離脱の受付・処理
 - 配送木情報の管理

- 配送木構築
- グループに属するユーザ情報の管理
- 配送木を構築するために必要な情報をユーザに通知

(3) 一般ユーザ

- グループへ参加の要求
- グループから離脱の要求
- 移動の通知
- 上流の中継ノードが居なくなった事の検知
- その他

3. 提案システムの設計

本章では提案システムの検討を行う。

3.1 システム設計の方針

マルチキャスト配送木の構成には ALM を用いて行う。これはモバイル端末などで無線を用いたユーザと有線を用いたユーザが混在する環境を想定しているためである。ネットワーク層でマルチキャスト配送木を構成するよりも実装の面で有利であり、また現状インターネット上のルータでマルチキャスト機能が有効化されていないことが多いためである。

(1) メンバ管理ノードが決めること

- マルチキャスト配送木
- どのユーザが中継ノードとなるか

(2) Shared Tree 方式を用いてマルチキャスト配送木を構築

- 配送木を変更しなければならない場合
- 配送木を変更した方が良い場合

(3) マルチキャスト配送木を構築する条件

- グループに参加するノードがあった時
- グループから離脱するノードがあった時
- ユーザから移動の通知があった時
- ユーザがいなくなったことを検知した時

(4) 一般ユーザが中継ノードになる

- 一つの中継ノードあたりの配送数の上限の設定
- 中継ノードがいなくなった時の対策
 - 中継ノードの下流ノードが上流のノードをいくつ把握するか

4. 提案システムの動作

本研究のシステムの特徴として、移動するユーザを考慮しオーバーレイネットワークでマルチキャスト配送木を構成するシステムがあげられる。本章ではシステムの動作の流れを述べる。

図4で示すように一般ユーザからの参加要求、離脱要求、移動の通知はメンバ管理ノードへ行われる。メンバ管理ノードはユーザからの要求によりユーザ情報の追加、削除、変更を行う。

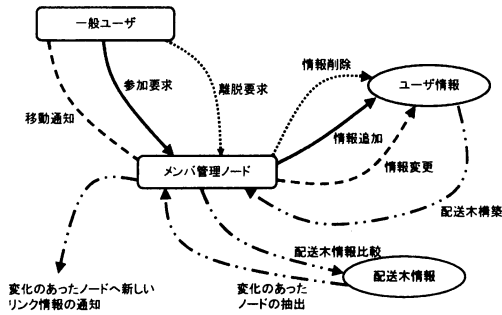


図4 動作の流れ

メンバ管理ノードは更新されたユーザ情報をもとに配送木構築を行い既存の配送木情報と比較する。比較により配送木に変化のあったノードを抽出し、配送木に変化のあったノードに対して新しいリンク情報を通知する。

中継ノードがダウンしたことの検知は中継ノードの上流ノードは一定回数データの送信に失敗した場合、下流ノードは一定時間データが配送されなかった場合に上流ノードがダウンしてないかを確認することで行う。

5. 今後の課題

提案手法をもとに今後実装を行い評価を行う。

● 評価項目

- (1) 提案手法で構築したマルチキャスト配送木の一般ユーザから見た
 - 他のユーザまでの平均ホップ数の平均
 - 最大ホップ数の平均
 を IP Multicast で構築した場合とでの比較
- (2) 配送木変更に掛るコスト
- (3) システム全体のバケット数、ノードが持つリンクあたりのバケット数

を評価項目とし、以下のそれぞれの場合について評価を行う

 - 中継ノードの配送するノード数を変化させる
 - ユーザの移動に伴い
 - (1) 配送木の変更を行わなければならない場合
 - (2) 配送木の変更を行う方が良い場合
 - (a) ユーザの移動が発生する度に配送木の再構築を行う場合
 - (b) ユーザの移動が発生してもある条件 (数, 移動距離) を満たすまで再構築を行わない場合

6. まとめ

本論文では移動するユーザを考慮した多地点間で多対多のグループコミュニケーションの提案を行った。また提案を実現するために必要な技術を述べ比較した。今後、実装と評価を行う。

- [1] Yang-hua Chu, Sanjay G. Rao and Hui Zhang, "A Case For End System Multicast", Proceedings of ACM SIGMETRICS, Santa Clara, CA, Jun. 2000.
- [2] D. Pendarakis, S. Shi, D. Verma and M. Waldvogel, "ALMI: An Application Level Multicast Infrastructure." Proceedings of the 3rd USNIX Symposium on Internet Technologies and Systems (USITS '01)
- [3] P. Francis, Yoid: Extending the Internet Multicast Architecture., <http://www.aciri.org/yoid/>, Apr. 2000.
- [4] J. Jannotti, D. K. Gifford, K. L. Johnson, M. F. Kaashoek, and J.W. O'Toole Jr., "Overcast: Reliable multicasting with an overlay network," Operating System Design and Implementation (OSDI), Oct. 2000.
- [5] Y. Chawathe. Scattercast: An Architecture for Internet Broadcast Distribution as an Infrastructure Service. Ph.D. Thesis, University of California, Berkeley, Dec. 2000.
- [6] Suman Banerjee, Bobby Bhattacharjee, and Christopher Kommareddy "Scalable Application Layer Multicast" in Proc. SIGCOMM, Aug. 2002.
- [7] Dave Kousiur, マスタリング TCP/IP マルチキャスト編, オーム社, 1999